DOCUMENT RESUME

ED 192 581 FL 011 741

AUTHOR Kaplan, Jeffrey P.
TITLE VP Anaphor Choice in Discourse.
PUB DATE 80
NOTE 12p.: Paper presented at the Annual Meeting of the University of Wisconsin-Milwaukee Linguistics Symposium (9th, Milwaukee, WI, March 7-8, 1980).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Discourse Analysis; Pronouns; *Sentence Structure; *Structural Grammar; *Verbs
IDENTIFIERS *Anaphora

ABSTRACT

Five types of Verb Phrase (VP) anaphors which appear to exist in free variation actually have specific grammatical, semantic, and rhetorical functions which account for the existence of some to the exclusion of others in certain syntactic environments. The five anaphors are: "do so," "do it," Verb Phrase Deletion (VPD), VPD with "too," and so-inversion. "Do so" and "do it" are the only true VP pro-forms. Certain grammatical restrictions on VPD with "too" and so-inversion that do not hold for the other three forms indicate that VPD with "too" and so-inversion are positional variants of each other. In natural speech, VPD with "too" and so-inversion are used only when subjects are compared; the others, elsewhere. Thus, where there was apparent free-variation, only more restricted forms occurred in a given environment. VPD with "too" and so-inversion are also divided along distributional and functional lines. "Too" usually occurs with one-word subjects, presents a marked, "new-old" information sequence, and an unmarked word order parallel to its preceding sentence. The function is to contrast subjects. "So" can occur with multiple subjects and produces a marked word order by subject-auxiliary inversion, thus highlighting information. Both syntactic structure and information sequencing act together, producing different rhetorical effects. (PJM)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
Jeffrey P. Kaplan
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED192581

FL011741

VP ANAPHOR CHOICE IN DISCOURSE

Jeff Kaplan
Northeastern University

0. Introduction. This paper is about factors that affect choice between types of VP anaphors which are grammatically possible in a given environment.* The anaphors to be discussed are listed in (1).

(1) a. do so: When Kennedy finishes his speech, he offers to shake hands with everyone in the audience, and for the next hour and fifteen minutes he does so. (Elizabeth Drew, The New Yorker[1])

b. do it: A: Say hi to Steve for me. B: I will do it for sure. (Natural conversation)

c. Verb Phrase Deletion (VPD): "Bruce, remember Tristam Shandy and that paper I copied from you?" "I certainly do Ø," said Gold with a touch of pique. (Joseph Heller, Good as Gold)

d. VPD with too: CW: I thought that was the ball.
JK: I did Ø too.
(Natural conversation)

e. so with inversion: Well, isn't Mass. unique? Don't we have crowded cities, antiquated roads, high accident rates? We do, indeed, but so do other Eastern states. (The Real Paper)

By "grammatically possible" I mean "judged acceptable by informants". The set of environments in which a VP anaphor is grammatically possible is generally larger than the set of environments in which it typically (or actually) occurs. For instance, do so and do it are grammatically possible in (2).

(2) John killed snakes and Bill did so/it too.

Examples like this abound in the linguistics literature. However, I have not found a single naturally occurring example of do so or do it in a structure like (2)--that is, in a coordinate conjunct, with a subject different from that of its antecedent, without an adverbial, and with too. I will try to give an explanation of this and similar apparent anomalies below.

*I am grateful to Nancy S. Levin, who provided much needed assistance, both substantive and editorial, during the gestation of this paper.
1. All naturally occurring data will be cited as to source. Made up examples and data from other linguists whose sources I don't know will have no citation.

1. Grammatical preface. First it will be necessary to present two of my assumptions from a sentence grammar analysis of verb phrase anaphors, since these assumptions are basic to my explanations of why certain grammatically possible anaphors do not occur in certain environments.

One assumption is that do so and do it are the only VP pro-forms on the agenda. The so of so aux NP (e.g., so do I) is not a pro-form. That it is not is supported by examples like (3), where there is no antecedent for it.

(3) JK: You have to remember, Carter ran a picture perfect campaign in New Hampshire.
EH: Well, yeah, but so is Bush running a good campaign. You watch out for him. He might actually win. (Natural conversation)

A second assumption is that what so of so did I is is a positional variant of the too in I did too. This is supported by a classical transformational argument involving shared restrictions. I will run through this argument briefly. 1. Both VPD with too and so with inversion must occur after their antecedents, never before. This is shown in (4).

(4) *After {everyone else had ∅ too / so had everyone else}, Harold grabbed a piece of birthday cake and gobbled it up.

Cf. OK: After everyone else had {∅ / done so}, ...

2. Both VPD with too and so with inversion must refer to a main VP, never to an embedded one. This is shown in (5).

(5) White Jade tried to catch the winner in the stretch. {So did Telferner ∅ / Telferner did ∅ too}.

In both versions, the reference can be only to tried to catch the winner in the stretch. In contrast, VPD or a pro-form refers to the embedded VP (Telferner DID (so)). 3. Neither VPD with too nor so with inversion can refer from a subordinate clause. This is shown in (6).

(6) *Henry ran a horse at Caliente before {John did ∅ too / so did John ∅}.

Cf. OK: ...before John did {∅ / so}.

4. Neither anaphor type can refer to an antecedent which is opposite it in polarity. This is shown in (7).

(7) a. *Wendy didn't buy any presents, but Sally did ∅ too/so did Sally.

Cf. OK: but Sally did {∅ / so}.

b. Wendy didn't buy any presents, but Sally didn't ∅ either/neither did Sally ∅.[2]

c. *Wendy bought a lot of presents, but Sally didn't ∅ either/neither did Sally ∅.

Cf. OK: But Sally didn't {∅ / do so}.

d. Wendy bought a lot of presents, but Sally did ∅ too/so did Sally ∅.

5. When the antecedent VP is in a preposed subordinate clause, both anaphor types are OK, except when the subordinate conjunction is unless, before, or until. This is shown in (8).

(8) {Because / Although / Since / When / *Unless / *Before / *Until} Sam was winning races, {Joe was ∅ too / so was Joe ∅}.[3]

Cf. OK: Before..., Joe was {∅ / doing so}.

These shared restrictions, and the synonymy of the constructions, support the claim that too and so are positional variants. The meaning of the element that shows up sometimes as one and sometimes as the other can be represented roughly as "what is predicated about the subject of one clause is also predicated about the subject of the other"[4]. (Although the justification for this meaning representation goes beyond intuition, its intuitive correctness can be seen in sentences like this one from Green 1968: Barb is seventeen, and WENDY is old enough to have a driver's license, too, which lacks VPD, and hence the crisscrossing predications can be seen clearly.)

These findings mean that the so with inversion construction is really a special case of Verb Phrase Deletion. Another implication is that the two so's are completely unrelated homonyms. I will show directly below how the typical environments in which VPD with so or too occurs naturally are distinct from those in which VPD without too or so, and the VP pro-forms, occur naturally. This is not surprising, in light of the distinctness of the so's. A third implication is that VPD with too and so with inversion are so close

2. Either and neither are complementary variants of too and so respectively which occur in negative contexts.

3. The unacceptable cases can be explained as violating the polarity restriction, since unless, before, and until all contain a negative element as part of their meaning.

4. This is based on a meaning representation for too proposed in Green 1968. For discussion, see Kaplan 1978.

grammatically that the possibility arises that they might be true free variants. In fact, it is difficult to show that they are not.

For convenience, I will refer to both VPD with so and VPD with too as SO/TOO VPD. "Naked" VPD, that is, without so or too, I will refer to as VPD. Do so and do it I will call PRO.

Let us now turn to factors affecting choice among the different anaphor types in environments in which more than one type can grammatically occur.

2. SO/TOO VPD vs. VPD and PRO. As can be seen from (4) - (8), the restrictions on SO/TOO VPD are not restrictions on VPD or PRO. What is more, the environments hospitable to SO/TOO VPD grammatically permit VPD and PRO as well. This is shown in (9).

(9) John befriended Harry; a. Bill should Ø. (VPD)

Bill should do so. (PRO)

b. so should Bill Ø. } (SO/TOO VPD)
Bill should Ø too. }

In other words, considering just the restrictions discussed so far, the domain of VPD and PRO is larger than, and properly includes, that of SO/TOO VPD. In the domain of SO/TOO VPD, we might expect free variation among SO/TOO VPD, VPD, and PRO. Data collected from naturally occurring speech, and from written sources, does not bear this out. In environments required by SO/TOO VPD, we find just that anaphor type, exclusively[5].

Before putting forward a hypothesis about why this is the case, it will be useful for me to discuss two characteristics of SO/TOO VPD environments that I have not yet mentioned.

First, SO/TOO VPD occurs typically as a complete second coordinate conjunct, or immediately following a sentence boundary, as, for example, in (10).

(10) a. (PK, playing "Mastermind":) Let's see, I've eliminated yellow.. yellow can go suck an egg...so can red and white Ø.. (natural conversation)

b. (Doctor to Bruce Gold:) Your weight is good and your heart and blood pressure are fine. Sid could lose some weight and use some exercise, but so could I Ø. (Good as Gold)

c. (Punch line of a story about getting into trouble in school:) He got his ass whipped and I did Ø too. (natural conversation)

d. (Watching football game on TV:) CW: I thought that was the ball.
JK: I did Ø too.
(natural conversation)

5. The fact that there are no cases of VPD or PRO in SO/TOO VPD environments may reflect a problem in the data, but that does not affect the point. We can simply take "0" as a case of a very low number.

Other structures are rare.[6]

Second, SO/TOO VPD always occurs with a subject different from that of its antecedent VP.[7] Observe that this holds for the examples in (10).

In part, a functional explanation is possible for why SO/TOO VPD requires a subject different from that of its antecedent. This has to do with a discourse constraint forbidding the repetition of strings which add no new information to the discourse. I will call this discourse constraint the Requirement for Proposition Differentiation.

(11) Requirement for Proposition Differentiation:

Do not repeat propositions without differentiating them.

The content of this constraint may be paraphrased roughtly as "Say something new"[8]. Now, in a typical SO/TOO VPD discourse, given that the VP is deleted and that SO/TOO VPD is the entire second clause, and given the grammatical restrictions on the construction summarized in sec. 1 (e.g., that the reference must be to a main VP and to an antecedent with the same polarity), the only constituents in the SO/TOO VPD clause that can differentiate its proposition from that of its antecedent clause are the aux and the subject. For reasons that are not clear, the aux's in both clauses are almost always identical. Therefore, in order to express a new proposition in the SO/TOO VPD clause, a speaker has to use a new subject.

Now let me make an initial speculation about why VPD and PRO do not occur in SO/TOO VPD's domain. The reason may have to do with a rather efficient functional specialization among the anaphor types, based on the requirement for proposition differentiation. Since PRO and VPD are not subject to the rigid constraints applying to SO/TOO VPD, many means of differentiating propositions, besides by having contrasting subjects, are available to them. Their propositions can be, and typically are, differentiated by adverbials, by having different illocutionary forces, by being embedded, by occurring on either side of or--all of which are unavailable, or nearly so, to SO/TOO VPD.[9] Ultimately because of the meaning of so/too, SO/TOO VPD just about has to

6. The rare structures include: a. Occurrence as an embedded VP: He is into fantasy a lot. And I've discovered so am I. b. Occurrence after a subordinate antecedent: When the season was over, so was the affair. (These are from The Homosexuals, by Alan Ebert.) c. Reference to a VP in a relative clause: I had been told about the New Englander, who found delight in his work. So did my Chicago acquaintance. d. Interrupted by a parenthetical phrase: I underestimated Miss Mason's spunkiness and her sense of self. So, apparently, did the PR department. (These are from Studs Terkel's Working.)

7. The morphemes so and too do not. They do not only when the second VP is not deleted because it is different from the first clause's VP. An example: Just as this hypothesis predicts that VPD covers a greater stretch of material to the RIGHT of a given V than do so/it, so does it predict that VPD can reach farther to the left of that V. (NL's dissertation)

8. See also Kaplan and Levin 1978.

9. SO/TOO VPD very rarely occurs after or. So/too's meaning contradicts that of exclusive or, though not that of inclusive or. Wherever so/too occurs after or, or is inclusive. In everyday speech, most occurrences of or are taken as being exclusive.

have its two related propositions differentiated by having different subjects. This difference perhaps is why VP anaphora discourses containing propositions differentiated by different subjects alone are more or less reserved for SO/TOO VPD.

This leads to the formulation of a general hypothesis: "All else being equal, given two or more forms which are grammatically permitted in an environment, one of which is permitted in only that environment, the more restricted form tends to be used in that environment to the exclusion of the less restricted form." Confirmation or disconfirmation of this hypothesis must await study of other cases of free variation in syntax.

3. SO vs. TOO. Unlike the choice between SO/TOO VPD, on the one hand, and PRO or VPD, on the other, the choice between PRO and VPD is largely a matter of grammaticality. That is, there does not seem to be an intersection of grammatical domains in which both constructions are possible and at least one occurs typically. Rather, their domains are largely distinct.[10] Consequently, in this paper I will ignore this issue, and turn instead to the question of what factors influence choice between NP aux Ø too, or TOO, and so aux NP Ø, or SO, the apparently freely varying versions of SO/TOO VPD.

I have observed two distributional differences between SO and TOO, one "external" and one "internal". Externally, there are many more cases of SO than there are of TOO. Well over 80% of my examples of SO/TOO VPD are cases of SO. Internally, the constructions differ in what kinds of subjects they take.

Let us look first at the difference in subjects. TOO occurs essentially only with a one-word subject, and the one word is almost always I. (When it is not, it is most frequently another pronoun.) SO occurs with a wide range of subject types. A simple pronoun is the most frequent type (and, again, I is the most frequent pronoun), but also found in respectable numbers are nouns, article + noun combinations, and much longer, "heavy", noun phrases.

Obviously the high percentage of pronoun subjects, and, in particular, cases of I, is a function of the use of these constructions in dialogues. In written English, predicate repetition is often effected otherwise than by SO/TOO VPD, e.g., by conjunction reduction (which is not possible with change of speaker).[11] This does not explain, however, the apparent greater freedom of SO to occur with many different types of subjects.

My informants report a difference in acceptability for SO and TOO when the anaphor's subject is very heavy. Consider (12).

(12) a. (From an editorial in Science 80 magazine:) By any measure, our social investment in science and technology and hence in change is enormous. So is our stake in understanding and managing that change.

b. ?By any measure, our social investment in science and technology and hence in change is enormous. Our stake in understanding and managing that change is Ø too.

10. Work in recent years by Ross, Kuno, Sag, Hankamer, Levin, and Kaplan has dealt with the differences between PRO and VPD.

11. I owe this point to Nancy S. Levin (personal communication).

The altered version, b, is judged lower in acceptability. This uncomfortability of heavy subjects with TOO is not surprising, in light of the general tendency to shift heavy NPs rightward; between SO and TOO only SO allows such a shift. This is a step toward explaining TOO's apparent preference for short subjects, but it does not explain why the subjects should so frequently be pronouns, and I in particular. The explanation for this may have something to do with the order in which old and new information is presented in SO and TOO, which rests on structural differences between the constructions. Examination of structure and of information sequencing may also explain different rhetorical effects which sometimes show up in these constructions.

Structurally, TOO has S-V order, which is the unmarked order for English. S-V order is also the usual order for SO/TOO VPD's antecedent clause. SO has subject-aux inversion, which is an unusual word order for English, and is never the order of the antecedent's clause.

The parallel structure between TOO and its antecedent makes high prominence possible in the contrast between the subjects of the clauses, and at the same time makes possible sentence-final emphatic stress on too. Example (13) exhibits a typical pattern.

(13) (JK is perusing the MLA job information list.)

JL: Hi! What are you doing?
JK: I'm looking for a job.
JL: I am Ø too. (natural conversation)

In the last line in (13), JL is communicating surprise. The contrastive pairing between the anaphor's subject and the antecedent's subject is reinforced by their occurring in the same position in their clauses.

SO seems less well suited for communicating surprise. Because its clause is not in parallel structure with its antecedent's clause, the contrast between its subject and its antecedent's subject is not so salient.

Corresponding to the word order differences between SO and TOO is a difference in the order in which information is presented. In TOO, information is presented in a marked way: the newest, most contrastive information, the subject, is presented first, flouting the Prague school dictum "old first, new last". In SO, the new information, the subject, is presented last.[12] In that respect, at least, then, SO is unmarked from the point of view of information sequencing. Observe, then, that TOO is unmarked from the point of view of usual word order, but marked from the point of view of information sequencing; SO is marked from the point of view of word order, but unmarked from the point of view of information sequencing. Possibly this is why SO is so much more common than TOO. Put otherwise, the fact that SO is so much more frequent in actual use than TOO may say something about the relative importance to the study of language use of syntactic structure and information structure.

This may also help explain why VPD and PRO do not occur in the environment for SO/TOO VPD, for VPD and PRO are just like TOO in their information order (as well as in their constituent order).

12. Technically, the deleted VP is last. But it is unaudible. The subject is the last tangible piece of information.

Now let us turn our attention to the fact that SO is an inversion construction. Inversion constructions have been studied by Green (1978), who identified a number of different inversion structures and a variety of rhetorical functions "distributed over, not among" them. One function Green identified was a "connective" function. Consider (14).

(14) Supporting merger were Lachlan Blair, Mary Blair, Phillips Garman, Linda Bronston, and David Lindstrom.
(=Green (7))

In a sentence like (14), according to Green, preposing the predicate performs "a connective function in indicating the relevance and importance of the individuals or entities to be named in the postponed (sic) subject to the events and issues that have been reported up to the point where they occur" (p. 6). However, in SO the predicate is not preposed, for, as we have seen, so is not a pro-form. Rather, it is a reflex of a complex operator (paraphrasable as something like "what is said about one subject is also said about the other"), and the actual predicate in a SO clause is the ∅ at the end.

Strictly speaking, there are two pieces of new information in SO/TOO VPD: the subject and so/too. However, the meaning of so/too contains a reference to the content of the previous clause. Thus it refers to old information, and, more or less implicitly, includes the old information in its meaning. In the form so, it occurs in the usual position for old information, clause-initially, and often does fulfill Green's "connective" function. Thus SO is unmarked from the point of view of information sequencing not only because it has new information last, but also because it has old information first. Frequently so also represents the "theme" (in Kuno's sense) of the discourse fragment composed of the pair of clauses. This is what seems to happen in (15) - (18).

(15) (TV commercial urging removal of architechtural barriers:)

Voice over: Who cares about architechtural barriers?
Chorus of replies by individuals in wheelchairs, with guide dogs, etc:

We do! I do!

Voice over: So should you.

(THEME: Caring about architechtural barriers.)

(16) HELEN (worried, toward Jack): But I can't just make love exactly when you want! It's impossible! I'm what I am and if that's not good enough for you then...(There was a half smile on her trembling lips.)

Dr. B: OK, you feel very deeply about this. And so does your husband here.
(Bach, The Intimate Enemy)

(THEME: Feeling very deeply about this.)

In (16), Dr. B. therapeutically intervenes, "changing the subject", that is, introducing a new theme.

(17) (Cains Mayonnaise commerical on TV:)

Hellman's has real vinegar! (New voice:) So does Cains.
Hellman's is made with real egg yolks! (New voice:) So is Cains.

(A few more turns)
.....
Hellman's has a real chemical additive! (New voice:) Cains's doesn't.

(THEMES: having real vinegar, having real egg yolks, etc.)

(18) (TV ad:)

The second shift is reporting. Some are handicapped. Viet Nam era veterans, economically disadvantaged...
Their employer gets a federal tax credit for hiring them, and so can you.

(THEME: Getting a federal tax credit for hiring them.)

TOO, on the other hand, is not suited to the connective function. Not only does it have new information first, it has old information (too) last. What TOO is suited for is a "contrastive" function, not only because it places contrasting elements in parallel slots, heightening the effect of contrastive stress, but also because it places new information in a marked position. In the case of both expressions, then, both syntactic structure and information sequence act together to produce rhetorical effects.

Let us return now to the discrepancy between SO and TOO in what subject types they occur with. If TOO emphasizes subject contrast by both position (new information first, instead of the unmarked last) and contrastive stress, heightened by parallel structure, wh should expect a wide range of subject types for TOO, especially subject types which can represent completely new, unexpected, contrastive information. However, my data does not support this. A large majority of cases with TOO have I as subject. Moreover, cases of SO that we have already discussed have contrastive stress on the postposed subject (e.g., (15) and (18)).

The explanation for this apparent anomaly may lie in the fact that although TOO's subject is contrastive, it is often not new information. I, the subject for TOO in almost all the cases I have collected, is not new information in a two-speaker dialogue. TOO, then, may be especially suited for a subject which represents old information but which the speaker wishes to highlight as contrastive.

The chart in (19) summarizes my findings with respect to SO and TOO.

(19)

	TOO	SO
Frequency	Low	High
Subjects	'Light' Pro I	Wide range (including 'heavy')
Rhetorical Function	Contrastive	Connective

4. Conclusion. I have discussed two cases of apparent free variation in syntax, the proper inclusion of the grammatical domain of SO/TOO VPD in that of VPD and PRO, and the near congruence of the grammatical domains of TOO and SO. In the latter case the free variation is real, in a syntactic sense; both forms do occur in many of the same syntactic environments. However, it proved possible to remove the free variation at a higher, discourse-functional level; the pair of constructions typically fulfill different functions in the discourse, because of both their different syntactic structures and the different orders in which the same information is presented in them. In the former case, we saw that the free variation was, in terms of actual use, essentially only apparent; no, or essentially no, cases of VPD or PRO are found in SO/TOO VPD's domain. Tentatively, I suggested that specialization of use of syntactic free variants exists; in a case of shared domain, a more restricted form takes precedence over a less restricted one. However, since from the point of view of information presentation PRO and VPD are the same as TOO, their actual non-occurrence in SO/TOO VPD's domain is probably due to the same cause as TOO's rarity in that domain. Thus in both cases what looks like a baffling syntactic puzzle receives a natural resolution framed in terms of information sequencing in discourse.

REFERENCES

Green, Georgia. 1968. On too and either, and not just too and either, either. Papers from the fourth regional meeting, Chicago Linguistic Society, ed. by B. Darden et. al., 22-39. Chicago: Chicago Linguistic Society.

_____. 1978. Discourse functions of inversion constructions. ms.

Hankamer, Jorge, and Ivan Sag. 1976. Deep and surface anaphora. Linguis-Inquiry 7.3, 391-426.

Kaplan, Jeff. 1978. A semantic theory of inverted so. Paper presented at the summer meeting, Linguistic Society of America, University of Illinois.

_____, and Nancy S. Levin. 1978. Say something new: a constraint on VP Deletion. Paper presented at the 53rd annual meeting, Linguistic Society of America, Boston, Massachusetts.

Kuno, Susumu. 1975. Conditions for Verb Phrase Deletion. Foundations of Language 13, 161-175.

Levin, Nancy S. 1979. Main verb ellipsis in spoken English. Ohio State University doctoral dissertation.

Ross, John R. 1974. Nearer to vee. Paper presented at LSA Linguistic Institute, Amherst, Massachusetts.

Sag, Ivan. 1976. Deletion and logical form. MIT doctoral dissertation.